



SEQUENCE LISTING

Amddt E
RECEIVED
JUL 31 2003
TECH CENTER 1600/2900

<110> Vaisvila, Romualdus
Morgan, Richard D.
Kucera, Rebecca B.
Claus, Toby B.
Raleigh, Elisabeth A.

<120> Method For Cloning And Producing The MseI Restriction
Endonuclease

<130> NEB-181

<140> US 09/689,343

<141> 2000-10-12

E7
<160> 21

<170> PatentIn version 3.1

<210> 1

<211> 903

<212> DNA

<213> Micrococcus sp.

<220>

<221> CDS

<222> (1)..(900)

<223>

<400> 1

atg cct atc tcg acc gtc tgg acg ccg gac gga gac gac ctc atc gtg 48
Met Pro Ile Ser Thr Val Trp Thr Pro Asp Gly Asp Asp Leu Ile Val
1 5 10 15

gag gcg gac aac ctc gat ttc att caa acg ctc ccc gac gcg agc ttc 96
Glu Ala Asp Asn Leu Asp Phe Ile Gln Thr Leu Pro Asp Ala Ser Phe
20 25 30

cga atg atc tac atc gat ccg ccg ttc aac aca ggg cga acg cag cgg 144
Arg Met Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Arg Thr Gln Arg
35 40 45

ctt cag tcg ctc aag acg acc cgc tcg gtc aca ggg tcg cga gtc ggc 192
Leu Gln Ser Leu Lys Thr Thr Arg Ser Val Thr Gly Ser Arg Val Gly
50 55 60

ttc aaa ggc cag acg tac gac acg gtc aag agc act ctg cac tcg tat 240
Phe Lys Gly Gln Thr Tyr Asp Thr Val Lys Ser Thr Leu His Ser Tyr
65 70 75 80

gac gac gct ttc acc gac tat tgg tcg ttc ctc gaa ccg cgt ctc ctg 288
Asp Asp Ala Phe Thr Asp Tyr Trp Ser Phe Leu Glu Pro Arg Leu Leu
85 90 95

gag gct tgg cgg ttg ctc acc cct gac ggc gcg ctc tat ctt cat ctg	336
Glu Ala Trp Arg Leu Leu Thr Pro Asp Gly Ala Leu Tyr Leu His Leu	
100 105 110	
gat tac cgc gag gtt cac tac gcc aag gtc gtc ctc gac gcg atg ttc	384
Asp Tyr Arg Glu Val His Tyr Ala Lys Val Val Leu Asp Ala Met Phe	
115 120 125	
gga cgc gaa agc ttc ctg aac gag ctg atc tgg gcg tac gac tac ggc	432
Gly Arg Glu Ser Phe Leu Asn Glu Leu Ile Trp Ala Tyr Asp Tyr Gly	
130 135 140	
gcg cgc tcg aag agc aag tgg ccc acc aag cac gac aac atc ctc gtg	480
Ala Arg Ser Lys Ser Lys Trp Pro Thr Lys His Asp Asn Ile Leu Val	
145 150 155 160	
tat gtg aag gac ccg aac aac tac gtc tgg aac ggt cag gat gta gat	528
Tyr Val Lys Asp Pro Asn Asn Tyr Val Trp Asn Gly Gln Asp Val Asp	
165 170 175	
cgc gag ccc tac atg gcg ccc ggg ctc gtt aca ccc gag aag gta gcg	576
Arg Glu Pro Tyr Met Ala Pro Gly Leu Val Thr Pro Glu Lys Val Ala	
180 185 190	
ctt ggc aag ctg ccc acc gac gtc tgg tgg cac aca atc gtt ccg cct	624
Leu Gly Lys Leu Pro Thr Asp Val Trp Trp His Thr Ile Val Pro Pro	
195 200 205	
gcg agc aaa gag cgc acc ggg tac gcg aca cag aag ccg gtc ggc atc	672
Ala Ser Lys Glu Arg Thr Gly Tyr Ala Thr Gln Lys Pro Val Gly Ile	
210 215 220	
atc cgt cgc atg att cag gcg agc agc aat gaa ggc gac tgg gtt ctg	720
Ile Arg Arg Met Ile Gln Ala Ser Ser Asn Glu Gly Asp Trp Val Leu	
225 230 235 240	
gat ttc ttc gct ggt agt ggg acg acc ggc gcc gcg gcc cgc cag ctc	768
Asp Phe Phe Ala Gly Ser Gly Thr Thr Gly Ala Ala Ala Arg Gln Leu	
245 250 255	
gga cgc cgt ttt gtg ctc gta gac gtc aac cca gaa gca atc gcg gta	816
Gly Arg Arg Phe Val Leu Val Asp Val Asn Pro Glu Ala Ile Ala Val	
260 265 270	
atg gca aaa cgg ttg gat gac ggg gca ttg gac acc agc gtg acg atc	864
Met Ala Lys Arg Leu Asp Asp Gly Ala Leu Asp Thr Ser Val Thr Ile	
275 280 285	
gtg cag act ccc cag agt gac cca cga acc gac gga tga	903
Val Gln Thr Pro Gln Ser Asp Pro Arg Thr Asp Gly	
290 295 300	

<210> 2

<211> 300

<212> PRT

<213> Micrococcus sp.

<400> 2

Met	Pro	Ile	Ser	Thr	Val	Trp	Thr	Pro	Asp	Gly	Asp	Asp	Leu	Ile	Val
1				5					10					15	
Glu	Ala	Asp	Asn	Leu	Asp	Phe	Ile	Gln	Thr	Leu	Pro	Asp	Ala	Ser	Phe
			20					25					30		
Arg	Met	Ile	Tyr	Ile	Asp	Pro	Pro	Phe	Asn	Thr	Gly	Arg	Thr	Gln	Arg
		35					40					45			
Leu	Gln	Ser	Leu	Lys	Thr	Thr	Arg	Ser	Val	Thr	Gly	Ser	Arg	Val	Gly
	50					55					60				
Phe	Lys	Gly	Gln	Thr	Tyr	Asp	Thr	Val	Lys	Ser	Thr	Leu	His	Ser	Tyr
65					70					75					80
Asp	Asp	Ala	Phe	Thr	Asp	Tyr	Trp	Ser	Phe	Leu	Glu	Pro	Arg	Leu	Leu
				85					90					95	
Glu	Ala	Trp	Arg	Leu	Leu	Thr	Pro	Asp	Gly	Ala	Leu	Tyr	Leu	His	Leu
			100					105					110		
Asp	Tyr	Arg	Glu	Val	His	Tyr	Ala	Lys	Val	Val	Leu	Asp	Ala	Met	Phe
		115					120					125			
Gly	Arg	Glu	Ser	Phe	Leu	Asn	Glu	Leu	Ile	Trp	Ala	Tyr	Asp	Tyr	Gly
	130					135					140				
Ala	Arg	Ser	Lys	Ser	Lys	Trp	Pro	Thr	Lys	His	Asp	Asn	Ile	Leu	Val
145					150					155					160
Tyr	Val	Lys	Asp	Pro	Asn	Asn	Tyr	Val	Trp	Asn	Gly	Gln	Asp	Val	Asp
				165					170					175	
Arg	Glu	Pro	Tyr	Met	Ala	Pro	Gly	Leu	Val	Thr	Pro	Glu	Lys	Val	Ala
			180					185					190		
Leu	Gly	Lys	Leu	Pro	Thr	Asp	Val	Trp	Trp	His	Thr	Ile	Val	Pro	Pro
		195					200					205			
Ala	Ser	Lys	Glu	Arg	Thr	Gly	Tyr	Ala	Thr	Gln	Lys	Pro	Val	Gly	Ile
	210					215					220				
Ile	Arg	Arg	Met	Ile	Gln	Ala	Ser	Ser	Asn	Glu	Gly	Asp	Trp	Val	Leu
225					230					235					240
Asp	Phe	Phe	Ala	Gly	Ser	Gly	Thr	Thr	Gly	Ala	Ala	Ala	Arg	Gln	Leu
				245					250					255	
Gly	Arg	Arg	Phe	Val	Leu	Val	Asp	Val	Asn	Pro	Glu	Ala	Ile	Ala	Val
			260					265					270		
Met	Ala	Lys	Arg	Leu	Asp	Asp	Gly	Ala	Leu	Asp	Thr	Ser	Val	Thr	Ile
		275					280					285			
Val	Gln	Thr	Pro	Gln	Ser	Asp	Pro	Arg	Thr	Asp	Gly				

```

290
295
300

<210> 3
<211> 1236
<212> DNA
<213> Unknown

<220>
<223> Environmental DNA

<220>
<221> CDS
<222> (1)..(1233)
<223>

<220>
<221> misc_feature
<222> (198)..(198)
<223> Xaa = any amino acid

<220>
<221> misc_feature
<222> (594)..(594)
<223> N= G, A, C or T

<400> 3

atg cct aca ctg gat tgg ccc ggt aaa cag tta agc ttc cca cca gct 48Met Pro
Thr Leu Asp Trp Pro Gly Lys Gln Leu Ser Phe Pro Pro Ala
1 5 10 15

acc tcc ttg cat ctg gag agt gtg gtc act gag gga gcg gag tca ccg 96
Thr Ser Leu His Leu Glu Ser Val Val Thr Glu Gly Ala Glu Ser Pro
20 25 30

cct aat cgt ctg att tgg gcg gac aac ctg ccg cta atg gta gat ttg 144
Pro Asn Arg Leu Ile Trp Ala Asp Asn Leu Pro Leu Met Val Asp Leu
35 40 45

ttg gcc gaa tat gaa ggg aaa atc gat ctg atc tac gcc gat ccc cct 192
Leu Ala Glu Tyr Glu Gly Lys Ile Asp Leu Ile Tyr Ala Asp Pro Pro
50 55 60

ttt ttt acg gat cgt act tat gcg gcg cga att ggt cat ggg gag gat 240
Phe Phe Thr Asp Arg Thr Tyr Ala Ala Arg Ile Gly His Gly Glu Asp
65 70 75 80

tcg cgt cgt cca caa acc tgg cag ctt gca gaa gga tat acg gac gag 288
Ser Arg Arg Pro Gln Thr Trp Gln Leu Ala Glu Gly Tyr Thr Asp Glu
85 90 95

tgg aag gat tta gat gaa tac ctg gac ttc ctt tat cca cgc ctg gta 336
Trp Lys Asp Leu Asp Glu Tyr Leu Asp Phe Leu Tyr Pro Arg Leu Val
100 105 110

ctg atg tat cga ctg ctg gca cca cac gga acg ctc tac ttg cac ctg 384

```

Leu	Met	Tyr	Arg	Leu	Leu	Ala	Pro	His	Gly	Thr	Leu	Tyr	Leu	His	Leu		
		115					120					125					
gac	tgg	cac	gcc	aat	gcc	tac	gta	cgt	gta	ctg	ctt	gat	gag	atc	ttc	432	
Asp	Trp	His	Ala	Asn	Ala	Tyr	Val	Arg	Val	Leu	Leu	Asp	Glu	Ile	Phe		
	130					135					140						
ggg	cga	cag	cgg	ttt	ctc	aac	gag	atc	gtc	tgg	atc	tat	cac	ggc	ccc	480	
Gly	Arg	Gln	Arg	Phe	Leu	Asn	Glu	Ile	Val	Trp	Ile	Tyr	His	Gly	Pro		
145					150					155					160		
tca	gcc	atc	cga	cgc	gcc	ttc	aag	cgc	aaa	cat	gat	acc	atc	ttg	gtt	528	
Ser	Ala	Ile	Arg	Arg	Ala	Phe	Lys	Arg	Lys	His	Asp	Thr	Ile	Leu	Val		
				165					170					175			
tat	gtg	aaa	ggc	gaa	aac	tat	aca	ttc	aat	gcg	gat	gcg	gtt	cgt	caa	576	
Tyr	Val	Lys	Gly	Glu	Asn	Tyr	Thr	Phe	Asn	Ala	Asp	Ala	Val	Arg	Gln		
			180					185					190				
cct	tac	cat	ccg	agc	acn	cat	aag	acc	ttc	gct	tcc	tcc	ccg	aag	gcc	624	
Pro	Tyr	His	Pro	Ser	Xaa	His	Lys	Thr	Phe	Ala	Ser	Ser	Pro	Lys	Ala		
		195					200					205					
ggc	ttt	ggc	aag	gtg	ccg	gat	ctg	cag	cgc	ggc	aaa	gtg	ccc	gaa	gac	672	
Gly	Phe	Gly	Lys	Val	Pro	Asp	Leu	Gln	Arg	Gly	Lys	Val	Pro	Glu	Asp		
	210					215				220							
tgg	tgg	tat	ttt	ccg	gtc	gtg	gcc	cgt	cta	cac	cga	gaa	cgg	agc	ggc	720	
Trp	Trp	Tyr	Phe	Pro	Val	Val	Ala	Arg	Leu	His	Arg	Glu	Arg	Ser	Gly		
225					230					235					240		
tat	ccg	act	caa	aag	cct	caa	gcc	ttg	ctg	gag	cgg	atc	ctg	ctg	gcc	768	
Tyr	Pro	Thr	Gln	Lys	Pro	Gln	Ala	Leu	Leu	Glu	Arg	Ile	Leu	Leu	Ala		
				245					250					255			
tcc	tcg	aac	gca	ggc	gat	ctg	gtg	gca	gac	ttc	ttc	tgc	ggc	tca	ggg	816	
Ser	Ser	Asn	Ala	Gly	Asp	Leu	Val	Ala	Asp	Phe	Phe	Cys	Gly	Ser	Gly		
			260					265					270				
aca	acc	gct	gtg	gtg	gca	gcc	cgt	ctg	gga	cgg	cgc	ttc	ctg	gtc	aac	864	
Thr	Thr	Ala	Val	Val	Ala	Ala	Arg	Leu	Gly	Arg	Arg	Phe	Leu	Val	Asn		
		275					280					285					
gat	gca	agc	tgg	cgc	gcc	gtt	cat	gtg	aca	cgc	aca	cgc	ttg	cta	cgc	912	
Asp	Ala	Ser	Trp	Arg	Ala	Val	His	Val	Thr	Arg	Thr	Arg	Leu	Leu	Arg		
	290					295					300						
gag	gga	gta	agt	ttc	act	ttt	gaa	cgc	cag	gaa	act	ttt	act	cta	cct	960	
Glu	Gly	Val	Ser	Phe	Thr	Phe	Glu	Arg	Gln	Glu	Thr	Phe	Thr	Leu	Pro		
305					310					315					320		
atc	cag	cca	ctt	cca	cca	gat	tgg	ttg	atc	atc	gcc	gag	gag	cag	att	1008	
Ile	Gln	Pro	Leu	Pro	Pro	Asp	Trp	Leu	Ile	Ile	Ala	Glu	Glu	Gln	Ile		
				325					330					335			
cgc	ctc	caa	gca	ccc	ttt	ctc	gta	gat	ttt	tgg	gaa	gtg	gac	gat	caa	1056	

Arg	Leu	Gln	Ala	Pro	Phe	Leu	Val	Asp	Phe	Trp	Glu	Val	Asp	Asp	Gln	
			340					345					350			
tgg	gat	ggc	aaa	atc	ttc	cgc	agc	cgt	cat	caa	ggc	tta	cgc	tcc	cgc	1104
Trp	Asp	Gly	Lys	Ile	Phe	Arg	Ser	Arg	His	Gln	Gly	Leu	Arg	Ser	Arg	
		355					360				365					
ctt	cag	gag	cag	gcg	ccg	ctc	tct	cta	cca	ttg	acc	ggg	aat	gga	ctg	1152
Leu	Gln	Glu	Gln	Ala	Pro	Leu	Ser	Leu	Pro	Leu	Thr	Gly	Asn	Gly	Leu	
	370					375					380					
ttg	tgt	gta	cgg	gta	gtg	agc	cgt	gaa	ggg	gaa	tac	tat	gag	ttc	aca	1200
Leu	Cys	Val	Arg	Val	Val	Ser	Arg	Glu	Gly	Glu	Tyr	Tyr	Glu	Phe	Thr	
385					390				395						400	
ggt	cga	gcc	gat	agc	cct	cac	ccc	gta	tcg	ttt	tga					1236
Gly	Arg	Ala	Asp	Ser	Pro	His	Pro	Val	Ser	Phe						
				405				410								

<210> 4
 <211> 411
 <212> PRT
 <213> Unknown

 <220>
 <223> Environmental DNA

 <220>
 <221> misc_feature
 <222> (198)..(198)
 <223> Xaa = any amino acid

 <400> 4

Met	Pro	Thr	Leu	Asp	Trp	Pro	Gly	Lys	Gln	Leu	Ser	Phe	Pro	Pro	Ala	
1				5				10					15			
Thr	Ser	Leu	His	Leu	Glu	Ser	Val	Val	Thr	Glu	Gly	Ala	Glu	Ser	Pro	
			20				25					30				
Pro	Asn	Arg	Leu	Ile	Trp	Ala	Asp	Asn	Leu	Pro	Leu	Met	Val	Asp	Leu	
		35				40					45					
Leu	Ala	Glu	Tyr	Glu	Gly	Lys	Ile	Asp	Leu	Ile	Tyr	Ala	Asp	Pro	Pro	
	50				55					60						
Phe	Phe	Thr	Asp	Arg	Thr	Tyr	Ala	Ala	Arg	Ile	Gly	His	Gly	Glu	Asp	
65				70					75					80		
Ser	Arg	Arg	Pro	Gln	Thr	Trp	Gln	Leu	Ala	Glu	Gly	Tyr	Thr	Asp	Glu	
			85				90						95			
Trp	Lys	Asp	Leu	Asp	Glu	Tyr	Leu	Asp	Phe	Leu	Tyr	Pro	Arg	Leu	Val	
		100					105					110				
Leu	Met	Tyr	Arg	Leu	Leu	Ala	Pro	His	Gly	Thr	Leu	Tyr	Leu	His	Leu	

115					120					125					
Asp	Trp	His	Ala	Asn	Ala	Tyr	Val	Arg	Val	Leu	Leu	Asp	Glu	Ile	Phe
130						135					140				
Gly	Arg	Gln	Arg	Phe	Leu	Asn	Glu	Ile	Val	Trp	Ile	Tyr	His	Gly	Pro
145					150					155					160
Ser	Ala	Ile	Arg	Arg	Ala	Phe	Lys	Arg	Lys	His	Asp	Thr	Ile	Leu	Val
				165					170					175	
Tyr	Val	Lys	Gly	Glu	Asn	Tyr	Thr	Phe	Asn	Ala	Asp	Ala	Val	Arg	Gln
			180					185					190		
Pro	Tyr	His	Pro	Ser	Xaa	His	Lys	Thr	Phe	Ala	Ser	Ser	Pro	Lys	Ala
		195					200					205			
Gly	Phe	Gly	Lys	Val	Pro	Asp	Leu	Gln	Arg	Gly	Lys	Val	Pro	Glu	Asp
	210					215					220				
Trp	Trp	Tyr	Phe	Pro	Val	Val	Ala	Arg	Leu	His	Arg	Glu	Arg	Ser	Gly
225					230					235					240
Tyr	Pro	Thr	Gln	Lys	Pro	Gln	Ala	Leu	Leu	Glu	Arg	Ile	Leu	Leu	Ala
				245					250					255	
Ser	Ser	Asn	Ala	Gly	Asp	Leu	Val	Ala	Asp	Phe	Phe	Cys	Gly	Ser	Gly
			260					265					270		
Thr	Thr	Ala	Val	Val	Ala	Ala	Arg	Leu	Gly	Arg	Arg	Phe	Leu	Val	Asn
		275					280					285			
Asp	Ala	Ser	Trp	Arg	Ala	Val	His	Val	Thr	Arg	Thr	Arg	Leu	Leu	Arg
	290					295					300				
Glu	Gly	Val	Ser	Phe	Thr	Phe	Glu	Arg	Gln	Glu	Thr	Phe	Thr	Leu	Pro
305					310					315					320
Ile	Gln	Pro	Leu	Pro	Pro	Asp	Trp	Leu	Ile	Ile	Ala	Glu	Glu	Gln	Ile
				325					330					335	
Arg	Leu	Gln	Ala	Pro	Phe	Leu	Val	Asp	Phe	Trp	Glu	Val	Asp	Asp	Gln
			340					345					350		
Trp	Asp	Gly	Lys	Ile	Phe	Arg	Ser	Arg	His	Gln	Gly	Leu	Arg	Ser	Arg
		355					360					365			
Leu	Gln	Glu	Gln	Ala	Pro	Leu	Ser	Leu	Pro	Leu	Thr	Gly	Asn	Gly	Leu
	370					375					380				
Leu	Cys	Val	Arg	Val	Val	Ser	Arg	Glu	Gly	Glu	Tyr	Tyr	Glu	Phe	Thr
385					390					395					400
Gly	Arg	Ala	Asp	Ser	Pro	His	Pro	Val	Ser	Phe					
				405					410						

<210> 5
 <211> 924
 <212> DNA
 <213> Unknown

<220>
 <223> Environmental DNA

<220>
 <221> CDS
 <222> (1)..(921)
 <223>

<400> 5

atg atc acg aac ctg atg gaa aac gat gtc att ggc aaa atc tac ttt	48
Met Ile Thr Asn Leu Met Glu Asn Asp Val Ile Gly Lys Ile Tyr Phe	
1 5 10 15	
gcc gac aac atg gaa gtc ctg cga ggg ctt ccg gcg gcg tcc gtg gac	96
Ala Asp Asn Met Glu Val Leu Arg Gly Leu Pro Ala Ala Ser Val Asp	
20 25 30	
ctg atc tac atc gat cct ccg ttc aac acc gga aag gtt cag gag cgc	144
Leu Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Lys Val Gln Glu Arg	
35 40 45	
act cag ctc aaa acg gtg cgc tcc gag tgg ggc gat cgc gtc gga ttc	192
Thr Gln Leu Lys Thr Val Arg Ser Glu Trp Gly Asp Arg Val Gly Phe	
50 55 60	
cag ggc cgt cgc tac gaa agc atc gtc gtg ggt aag aag cgc ttt acc	240
Gln Gly Arg Arg Tyr Glu Ser Ile Val Val Gly Lys Lys Arg Phe Thr	
65 70 75 80	
gac ttc ttc gac gac tat ctg gct ttc ctg gaa ccg cgc ctg gtc gaa	288
Asp Phe Phe Asp Asp Tyr Leu Ala Phe Leu Glu Pro Arg Leu Val Glu	
85 90 95	
gcc cat cgt gtt ctg gcg ccg cac ggg tgc ctc tac ttt cac gtc gac	336
Ala His Arg Val Leu Ala Pro His Gly Cys Leu Tyr Phe His Val Asp	
100 105 110	
tac cgc gag gtg cac tac tgt aag gtc ctt ctt gac ggc atc ttc ggt	384
Tyr Arg Glu Val His Tyr Cys Lys Val Leu Leu Asp Gly Ile Phe Gly	
115 120 125	
cgc gag gcc ttt ctc aac gag atc atc tgg gcc tac gat tac ggc ggg	432
Arg Glu Ala Phe Leu Asn Glu Ile Ile Trp Ala Tyr Asp Tyr Gly Gly	
130 135 140	
cgt ccg aag gac agg tgg cct cct aag cac gac aac atc ctg ctc tac	480
Arg Pro Lys Asp Arg Trp Pro Pro Lys His Asp Asn Ile Leu Leu Tyr	
145 150 155 160	
gcc aag act ccc ggt cgc cac gtg ttc aat gcg gac gaa atc gag cgc	528
Ala Lys Thr Pro Gly Arg His Val Phe Asn Ala Asp Glu Ile Glu Arg	
165 170 175	

att ccc tac atg gct ccg ggc ctg gtt ggc ccc gaa aag gca gcc cgt 576
 Ile Pro Tyr Met Ala Pro Gly Leu Val Gly Pro Glu Lys Ala Ala Arg
 180 185 190

gga aaa ctg cca acc gac acg tgg tgg cat acg atc gtt ccg acc agc 624
 Gly Lys Leu Pro Thr Asp Thr Trp Trp His Thr Ile Val Pro Thr Ser
 195 200 205

ggc tcc gag aag acc ggg tat cca acc cag aaa cct tta ggg att ctc 672
 Gly Ser Glu Lys Thr Gly Tyr Pro Thr Gln Lys Pro Leu Gly Ile Leu
 210 215 220

cgc cgt att gtg cag gca tcg tct cat ccg ggg gca gtc gtg ctc gac 720
 Arg Arg Ile Val Gln Ala Ser Ser His Pro Gly Ala Val Val Leu Asp
 225 230 235 240

ttc ttc gcc ggc agt ggg aca aca ggg gta gcg gct ttt gag ttg ggc 768
 Phe Phe Ala Gly Ser Gly Thr Thr Gly Val Ala Ala Phe Glu Leu Gly
 245 250 255

cgg cgt ttc att ctg gtc gat aac cat ccg gag gcc ctc cag gtg atg 816
 Arg Arg Phe Ile Leu Val Asp Asn His Pro Glu Ala Leu Gln Val Met
 260 265 270

gcc agg cgc ttc gac ggc atc gag ggg atc gaa tgg gtg ggc ttc gat 864
 Ala Arg Arg Phe Asp Gly Ile Glu Gly Ile Glu Trp Val Gly Phe Asp
 275 280 285

ccg aca ccg tac cag aag ggc gca aag cag cgc cgc tcc tgc ccg gcg 912
 Pro Thr Pro Tyr Gln Lys Gly Ala Lys Gln Arg Arg Ser Cys Pro Ala
 290 295 300

ccc acc ggg taa 924
 Pro Thr Gly
 305

<210> 6
 <211> 307
 <212> PRT
 <213> Unknown

<220>
 <223> Environmental DNA

<400> 6

Met Ile Thr Asn Leu Met Glu Asn Asp Val Ile Gly Lys Ile Tyr Phe
 1 5 10 15

Ala Asp Asn Met Glu Val Leu Arg Gly Leu Pro Ala Ala Ser Val Asp
 20 25 30

Leu Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Lys Val Gln Glu Arg
 35 40 45

Thr Gln Leu Lys Thr Val Arg Ser Glu Trp Gly Asp Arg Val Gly Phe
 50 55 60
 Gln Gly Arg Arg Tyr Glu Ser Ile Val Val Gly Lys Lys Arg Phe Thr
 65 70 75 80
 Asp Phe Phe Asp Asp Tyr Leu Ala Phe Leu Glu Pro Arg Leu Val Glu
 85 90 95
 Ala His Arg Val Leu Ala Pro His Gly Cys Leu Tyr Phe His Val Asp
 100 105 110
 Tyr Arg Glu Val His Tyr Cys Lys Val Leu Leu Asp Gly Ile Phe Gly
 115 120 125
 Arg Glu Ala Phe Leu Asn Glu Ile Ile Trp Ala Tyr Asp Tyr Gly Gly
 130 135 140
 Arg Pro Lys Asp Arg Trp Pro Pro Lys His Asp Asn Ile Leu Leu Tyr
 145 150 155 160
 Ala Lys Thr Pro Gly Arg His Val Phe Asn Ala Asp Glu Ile Glu Arg
 165 170 175
 Ile Pro Tyr Met Ala Pro Gly Leu Val Gly Pro Glu Lys Ala Ala Arg
 180 185 190
 Gly Lys Leu Pro Thr Asp Thr Trp Trp His Thr Ile Val Pro Thr Ser
 195 200 205
 Gly Ser Glu Lys Thr Gly Tyr Pro Thr Gln Lys Pro Leu Gly Ile Leu
 210 215 220
 Arg Arg Ile Val Gln Ala Ser Ser His Pro Gly Ala Val Val Leu Asp
 225 230 235 240
 Phe Phe Ala Gly Ser Gly Thr Thr Gly Val Ala Ala Phe Glu Leu Gly
 245 250 255
 Arg Arg Phe Ile Leu Val Asp Asn His Pro Glu Ala Leu Gln Val Met
 260 265 270
 Ala Arg Arg Phe Asp Gly Ile Glu Gly Ile Glu Trp Val Gly Phe Asp
 275 280 285
 Pro Thr Pro Tyr Gln Lys Gly Ala Lys Gln Arg Arg Ser Cys Pro Ala
 290 295 300
 Pro Thr Gly
 305

<210> 7

<211> 561

<212> DNA

<213> Micrococcus sp.

<220>

<221> CDS

<222> (1)..(558)

<223>

<400> 7

gtg	acc	cac	gaa	ccg	acg	gat	gat	ccc	gat	ttc	ata	gtg	atg	gcc	gcg	48
Val	Thr	His	Glu	Pro	Thr	Asp	Asp	Pro	Asp	Phe	Ile	Val	Met	Ala	Ala	
1			5					10						15		

agc	gcg	gcg	aac	ctc	gct	gat	cgg	tac	gta	gcg	agt	gaa	gac	gac	ccc	96
Ser	Ala	Ala	Asn	Leu	Ala	Asp	Arg	Tyr	Val	Ala	Ser	Glu	Asp	Asp	Pro	
			20					25					30			

tgg	gtc	ggc	agc	ccg	ttc	gag	tgg	atc	ctt	cgc	gtt	cca	tcc	aga	acg	144
Trp	Val	Gly	Ser	Pro	Phe	Glu	Trp	Ile	Leu	Arg	Val	Pro	Ser	Arg	Thr	
		35					40					45				

aag	ggc	gcg	gtc	ggg	gag	ctg	ctc	gtg	agc	gaa	tgg	gct	aat	gcc	aaa	192
Lys	Gly	Ala	Val	Gly	Glu	Leu	Leu	Val	Ser	Glu	Trp	Ala	Asn	Ala	Lys	
	50					55					60					

ggc	ctc	cgt	gtg	aag	agg	tcg	ggg	tcc	agc	gat	gcg	gac	cgc	gtg	atc	240
Gly	Leu	Arg	Val	Lys	Arg	Ser	Gly	Ser	Ser	Asp	Ala	Asp	Arg	Val	Ile	
65				70						75					80	

aac	ggg	cat	cgc	atc	gag	atc	aag	atg	tcg	act	ttg	tgg	aag	tcc	ggc	288
Asn	Gly	His	Arg	Ile	Glu	Ile	Lys	Met	Ser	Thr	Leu	Trp	Lys	Ser	Gly	
			85						90					95		

ggc	ttc	aag	ttt	cag	cag	atc	cgg	gat	cag	gag	tac	gac	ttt	tgc	ctc	336
Gly	Phe	Lys	Phe	Gln	Gln	Ile	Arg	Asp	Gln	Glu	Tyr	Asp	Phe	Cys	Leu	
			100					105					110			

tgc	ctt	ggg	atc	agc	ccg	ttc	gaa	gtg	cac	gcg	tgg	ctg	ctg	ccc	aaa	384
Cys	Leu	Gly	Ile	Ser	Pro	Phe	Glu	Val	His	Ala	Trp	Leu	Leu	Pro	Lys	
		115					120					125				

gac	cta	ttg	ctt	gag	tac	gtg	att	ggg	cac	atg	ggg	cag	cac	acc	ggc	432
Asp	Leu	Leu	Leu	Glu	Tyr	Val	Ile	Gly	His	Met	Gly	Gln	His	Thr	Gly	
	130					135					140					

gcg	agc	ggg	agc	gac	act	gcg	tgg	ctg	ggg	ttc	cca	gcg	gac	gag	ccg	480
Ala	Ser	Gly	Ser	Asp	Thr	Ala	Trp	Leu	Gly	Phe	Pro	Ala	Asp	Glu	Pro	
145					150					155					160	

tat	gac	tgg	atg	cgc	cct	ttc	gga	ggg	cgc	tta	ggg	cac	gtc	gaa	gat	528
Tyr	Asp	Trp	Met	Arg	Pro	Phe	Gly	Gly	Arg	Leu	Gly	His	Val	Glu	Asp	
				165					170					175		

ctc	ctc	ctc	gcg	gcc	ggc	ccc	ggg	ccc	tac	tga						561
Leu	Leu	Leu	Gla	Ala	Gly	Pro	Gly	Pro	Tyr							
			180					185								

<210> 8
 <211> 186
 <212> PRT
 <213> Micrococcus sp.

<400> 8

Val	Thr	His	Glu	Pro	Thr	Asp	Asp	Pro	Asp	Phe	Ile	Val	Met	Ala	Ala	
1				5					10					15		
Ser	Ala	Ala	Asn	Leu	Ala	Asp	Arg	Tyr	Val	Ala	Ser	Glu	Asp	Asp	Pro	
			20					25					30			
Trp	Val	Gly	Ser	Pro	Phe	Glu	Trp	Ile	Leu	Arg	Val	Pro	Ser	Arg	Thr	
		35					40					45				
Lys	Gly	Ala	Val	Gly	Glu	Leu	Leu	Val	Ser	Glu	Trp	Ala	Asn	Ala	Lys	
	50					55					60					
Gly	Leu	Arg	Val	Lys	Arg	Ser	Gly	Ser	Ser	Asp	Ala	Asp	Arg	Val	Ile	
65					70					75					80	
Asn	Gly	His	Arg	Ile	Glu	Ile	Lys	Met	Ser	Thr	Leu	Trp	Lys	Ser	Gly	
				85					90					95		
Gly	Phe	Lys	Phe	Gln	Gln	Ile	Arg	Asp	Gln	Glu	Tyr	Asp	Phe	Cys	Leu	
			100					105					110			
Cys	Leu	Gly	Ile	Ser	Pro	Phe	Glu	Val	His	Ala	Trp	Leu	Leu	Pro	Lys	
		115					120					125				
Asp	Leu	Leu	Leu	Glu	Tyr	Val	Ile	Gly	His	Met	Gly	Gln	His	Thr	Gly	
	130					135					140					
Ala	Ser	Gly	Ser	Asp	Thr	Ala	Trp	Leu	Gly	Phe	Pro	Ala	Asp	Glu	Pro	
145					150					155					160	
Tyr	Asp	Trp	Met	Arg	Pro	Phe	Gly	Gly	Arg	Leu	Gly	His	Val	Glu	Asp	
				165					170					175		
Leu	Leu	Leu	Ala	Ala	Gly	Pro	Gly	Pro	Tyr							
			180					185								

<210> 9
 <211> 413
 <212> DNA
 <213> Escherichia coli

<400> 9

accggtgatt	ggacattgcc	gaaatcaggc	tgtctctcac	tatttgacgc	actggctgga	60
ctatccacat	ctaccttatt	ccccgaata	acgagatccc	ttccagcacc	gggcaattgc	120
ccggtttttt	ttgcgttgaa	tttgtcattt	tgtgccgtgg	tgtttaaacc	gcacagaata	180
aattgtcgtg	atttcacctt	taaaataaaa	ttaaaagaga	aaaaaattct	ctgtggaagg	240

gctatgtag ataaaattga ccgtaagctg ctggccttac tgcagcagga ttgcaccctc 300
tctttgcagg cactgggtga agccgttaat ctgacaacca ccccttgctg gaagcgctg 360
aaacggctgg aggacgacgg tacccttacc ggcaaagtcg ccctgctgga tcc 413

<210> 10
<211> 25
<212> PRT
<213> Micrococcus sp.

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> Xaa = any amino acid

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> Xaa = any amino acid

<400> 10

Met Thr His Glu Pro Thr Asp Asp Pro Asp Phe Ile Val Met Ala Ala
1 5 10 15

Ser Ala Xaa Asn Leu Ala Asp Xaa Tyr
20 25

<210> 11
<211> 28
<212> DNA
<213> unknown<220><223> Primer MseI-IP1

<400> 11
cttctgcagc cgatttcata gtgatggc 28

<210> 12
<211> 27
<212> DNA
<213> Unknown

<220>
<223> Primer MseI-IP2

<400> 12

gttctgcaga tcgggatcat ccgtcgg 27

<210> 13
<211> 49
<212> DNA

<213> unknown

<220>

<223> Primer MseI-IP3

<400> 13

ggttctgcag ttaaggaggt ttaacatatg acccacgaac cgacggatg

49

<210> 14

<211> 35

<212> DNA

<213> unknown

<220>

<223> Primer MseI-IP4

<400> 14

gttggatccg tcgacgcttc tcggcgtacc gagcg

35

<210> 15

<211> 32

<212> DNA

<213> unknown

<220>

<223> Primer

<400> 15

gaaccgatc cgaccctgag tgagaacatg cc

32

<210> 16

<211> 36

<212> DNA

<213> Unknown

<220>

<223> primer

<400> 16

aggtcgcaat tgccaggggt cgtcttcact cgctac

36

<210> 17

<211> 30

<212> DNA

<213> unknown

<220>

<223> primer

<400> 17

gcgatacaga ccggttcaga caggataaaag

30

<210> 18

<211> 24

<212> DNA

<213> unknown

<220>

<223> primer

<400> 18

ggtcgggatcc ggcgatacag cgag

24

<210> 19

<211> 24

<212> DNA

<213> unknown

<220>

<223> primer

<400> 19

ggatcttcca gtggtgcatg aacg

24

<210> 20

<211> 34

<212> DNA

<213> unknown

<220>

<223> forward primer

<400> 20

agactcccc atatgaccca cgaaccgacg gatg

34

<210> 21

<211> 34

<212> DNA

<213> unknown

<220>

<223> reverse primer

<400> 21

gggtggtccc gctagctatt agtagggacc gggg

34

En
conc ✓